

# Revisiting the evolutionary history of domestic and wild ducks based on genomic analyses

Xing Guo<sup>1,#</sup>, Xin-Xin He<sup>1,#</sup>, Hong Chen<sup>1,#</sup>, Zhi-Cheng Wang<sup>1</sup>, Hui-Fang Li<sup>2</sup>, Jiang-Xian Wang<sup>1</sup>, Ming-Shan Wang<sup>3,4,\*</sup>, Run-Shen Jiang<sup>1,\*</sup>

<sup>1</sup> Anhui Province Key Laboratory of Local Livestock and Poultry Genetic Resource Conservation and Bio-breeding, Anhui Agricultural University, Hefei, Anhui 230036, China

<sup>2</sup> Jiangsu Institute of Poultry Science, Chinese Academy of Agriculture Science, Yangzhou, Jiangsu 225125, China

<sup>3</sup> Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, CA 95064, USA

<sup>4</sup> Department of Ecology and Evolutionary Biology, University of California Santa Cruz, CA 95064, USA

## ABSTRACT

Although domestic ducks have been important poultry species throughout human history, their origin remains enigmatic, with mallards and/or Chinese spot-billed ducks being proposed as the direct wild ancestor(s) of domestic ducks. Here, we analyzed 118 whole genomes from mallard, Chinese spot-billed, and domestic ducks to reconstruct their evolutionary history. We found pervasive introgression patterns among these duck populations. Furthermore, we showed that domestic ducks separated from mallard and Chinese spot-billed ducks nearly 38 thousand years ago (kya) and 54 kya, respectively, which is considerably outside the time period of presumed duck domestication. Thus, our results suggest that domestic ducks may have originated from another wild duck population that is currently undefined or unsampled, rather than from present-day mallard and/or Chinese spot-billed ducks, as previously thought. Overall, this study provides new insight into the complex evolution of ducks.

## Open Access

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2021 Editorial Office of Zoological Research, Kunming Institute of Zoology, Chinese Academy of Sciences

**Keywords:** Duck; Population structure; Demographic histories; Domestication

## INTRODUCTION

Ducks are among the most economically important waterfowl species in the world and provide a source of high-quality protein for many people (Huang et al., 2013). China is thought to be one of the major centers of duck domestication (Hitosugi et al., 2007; Larson et al., 2014). Based on archaeological evidence, the domestication of ducks took place approximately 1 000–2 500 years ago (Museum, 1979). However, there are 14 taxa recognized in the mallard complex (Lavretsky et al., 2014), among which mallards and spot-billed ducks, two ecologically important and geographically widespread species, are more closely related than any other wild species to modern domestic ducks (Hou et al., 2012; Li et al., 2010). A recent study on mtDNA D-loop sequence variation among 50 duck breeds from eight countries suggested that domestic ducks are descended from wild mallards (Hou et al., 2012). Furthermore, based on 12 microsatellite markers of 958 ducks, Zhang et al. (2014) also suggested a single origin for the domestic duck from wild

Received: 23 September 2020; Accepted: 30 November 2020; Online: 03 December 2020

Foundation items: This study was supported by the Natural Science Foundation of Anhui Province (1908085QC143) and the Natural Science Foundation for Young Scholars of Anhui Agricultural University (yj2018-51)

\*Authors contributed equally to this work

\*Corresponding authors, E-mail: mwang153@ucsc.edu; jiangrunshen@ahau.edu.cn

DOI: 10.24272/j.issn.2095-8137.2020.133

mallards. However, comparisons of polymorphisms within 667 bp mtDNA D-loop sequences from 26 Chinese domestic duck breeds and two wild duck species showed that both mallards and Chinese spot-billed ducks are wild ancestors of domestic ducks (Li et al., 2010). This result is supported by analysis of mtDNA data from Korean ducks (Jin et al., 2014). However, despite considerable effort, the origin and the evolutionary history of domestic ducks are still under debate.

The main factor confounding our understanding of these issues is that the evolutionary history of wild ducks, particularly that of mallards and Chinese spot-billed ducks, is incomplete. Mallards and Chinese spot-billed ducks are phenotypically distinct (Wang et al., 2019). However, recent studies have shown that these two wild duck species are only weakly genetically differentiated. For instance, based on data from 20 nuclear loci, Lavretsky et al. (2014) found that Chinese spot-billed ducks are indistinguishable from Eurasia mallards ( $\Phi_{ST}=0$ ), in agreement with Wang et al. (2019), who reported low levels of divergence for the two species with average  $F_{ST}$  values of 0.014. Moreover, mallards and Chinese spot-billed ducks can hybridize in nature (Kulikova et al., 2003; Wang et al., 2019), which complicates the evolutionary histories of these wild species and domestic ducks.

Most previous research on duck evolutionary history is based on mtDNA and short nuclear segments. Due to its uniparental inheritance, mtDNA is sensitive to genetic drift/bottlenecks, sex-biased selection, and admixture, and therefore shows low efficiency in the reconstruction of complete evolutionary histories, especially with respect to domestication (Larson & Burger, 2013; Park et al., 2015; Wang et al., 2014; Zhang et al., 2020). Additionally, short nuclear segments provide limited information and low resolution in clarifying the complex evolutionary history of species, especially for population pairs with weak differentiation. In recent years, the reference genomes of many domestic species (e.g., chickens, ducks, dogs, goats, pigs, horses) have been published (Dong et al., 2013; Groenen et al., 2012; Hillier et al., 2004; Huang et al., 2013; Lindblad-Toh et al., 2005; Orlando et al., 2013). With the rapid development of whole-genome sequencing technologies and considerable reductions in the costs of sequencing, re-sequencing the genomes of individuals at the population scale has become a reality. Analyses based on genome-wide data have revolutionized research on the evolutionary histories of domestic animals. For example, based on mtDNA data, chickens are thought to have undergone multiple, independent domestications in South and Southeast Asia (Liu et al., 2006; Miao et al., 2013). However, large-scale whole-genome data have also revealed that domestic chickens are derived from *Gallus gallus spadiceus*, which is predominantly distributed in southwestern China, northern Thailand, and Myanmar (Wang et al., 2020). A similar case has also been reported in domestic dogs, with mtDNA analysis inconsistent with an East Asian origin 15 thousand years ago (kya) (Savolainen et al. 2002) or European domestication 18.8 kya to 32.1 kya (Thalmann et al., 2013). Furthermore, whole-genome data

suggest that dogs originated from southern East Asian wolves 33 kya (Wang et al., 2016). Whole-genome data analyses have uncovered the genetic basis underlying remarkable behavioral, physiological, and morphological traits, and have been applied in many domestic species (Guo et al., 2016; Plassais et al., 2019; Wang et al., 2016). Whole-genome re-sequencing has also been applied in duck genomic research. For instance, Zhang et al. (2018) performed comparative population genomics between domestic ducks and mallards and identified several positive selection genes that affect brain and neuronal development in domestic ducks. By sequencing 1132 genomes, Zhou et al. (2018) performed genome-wide association study (GWAS) and selective sweeps analysis and revealed that the *MITF* and *IGF2BP1* genes are associated with plumage color and body size in ducks. However, the aforementioned studies mainly focused on the genetic basis underlying specific traits, and our understanding of the evolutionary history of mallards, Chinese spot-billed ducks, and domestic ducks at the genome-wide scale remains limited. In this study, we analyzed 118 whole genomes of mallards, Chinese spot-billed ducks, and domestic ducks to investigate the genetic structure and evolutionary history of the two wild species and to determine whether modern ducks originated from mallards and/or Chinese spot-billed ducks.

## MATERIALS AND METHODS

### Samples, read mapping, and single-nucleotide polymorphism (SNP) calling

Domestic ducks are generally bred for egg laying, meat production, or both (dual-purpose) (Zhang et al., 2018). Recent molecular data have indicated that meat, egg, and dual purpose-type ducks were domesticated from mallards and/or Chinese spot-billed ducks (Li et al., 2010). To reconstruct the evolutionary history of these duck populations, we downloaded the genomic data of 118 ducks from the EBI (<https://www.ebi.ac.uk/ena>) and GSA databases (<http://bigd.big.ac.cn/gsa/>), including one Muscovy duck, eight Chinese spot-billed ducks (eastern spot-billed duck), 25 mallards, 38 meat-type ducks (MET) (8 Cherry valley, 22 Pekin, and 8 maple leaf ducks), and 46 egg and dual purpose-type ducks (EDT) (11 Jinding, 8 Shanma, 16 Shaoxing, and 11 Gaoyou ducks) (Zhang et al., 2018; Zhou et al., 2018). The geographic distribution of these samples is provided in Supplementary Figure S1 and sample information is described in Supplementary Table S1.

Low-quality bases with Phred-quality scores <20 were discarded using Btrim with parameters "-s -a 20 -q" (Kong, 2011). The BWA-MEM algorithm (Li, 2014) was then used to map the high-quality reads to the duck reference genome (Zhou et al., 2018) with default settings, except the "-t 6 -M -R <\$readgroup>" options. The bam alignment files were sorted, and multiply aligned reads were removed. SNP calling was performed using UnifiedGenotyper with default parameters, and filtering was performed with VariantFiltration tools ("QUAL<40.0 MQ<25.0 MQ0≥4 & ((MQ0/(1.0\*DP))>0.1)

-cluster 3 -window 10") in GATK (McKenna et al., 2010).

### Population diversity, structure, and relationships

We used VCFtools v0.1.13 (Danecek et al., 2011) with non-overlapping 50 kb windows (--window-pi 50000) to estimate nucleotide diversity ( $\pi$ ) for each domestic breed and the two wild populations. Linkage disequilibrium (LD) was estimated using PLINK (Chang et al., 2015) based on eight individuals from each duck population. We only used SNPs with minor allele frequencies (MAF) greater than 0.05. To explore the relationships between the different duck populations, principal component analysis (PCA) was performed based on autosomal SNPs using genome-wide complex trait analysis (GCTA) (Yang et al., 2011). The SNPs were pruned using PLINK with the settings "--indep-pairwise 50 10 0.1". The genetic relationship matrix was calculated with the "--make-grm" option, and then three principal components were estimated with the "--pca3" option. To reveal the phylogenetic relationships of all ducks, MEGA7 (Kumar et al., 2016) and FastTree (Price et al., 2010) were employed to construct neighbor-joining (NJ) and maximum-likelihood (ML) trees with 100 bootstraps. We also used VCFtools v0.1.13 (Danecek et al., 2011) with non-overlapping 50 kb windows (--fst-window-size 50000) to estimate the  $F_{ST}$  between all possible pairwise combinations. ADMIXTURE was used to estimate the population structure of the two wild duck species ( $K$  was assumed to range from 2 to 10), followed by the addition of the domestic ducks with  $K$  values ranging from 2 to 11 (Alexander et al., 2009). The estimated ancestry for each cluster was visualized using the ggplot2 R package.

### Introgression analysis

D-statistics (Patterson et al., 2012) were used to verify the existence of gene flow between wild and domestic ducks and between mallards and Chinese spot-billed ducks. D-statistics are also known as the ABBA-BABA test. Given a topology (O, (P3, (P1, P2))), the ABBA pattern reflects gene flow between P3 and P1, whereas BABA events reflect gene flow between P3 and P2. Significance is expressed as Z-scores, which are generally considered significant when their absolute values are higher than 3. To determine the gene flow between mallards and Chinese spot-billed ducks, we used Muscovy ducks as the outgroup (O) and mallard, Chinese spot-billed, and domestic duck breeds as the P3, P1, and P2 groups, respectively. A negative  $D$  value with  $Z$ -score  $< -3$  indicates admixture between mallards and Chinese spot-billed ducks, whereas a positive  $D$  value with  $Z$ -score  $> 3$  indicates admixture between mallards and domestic ducks. We also estimated the genetic introgression between Chinese spot-billed and domestic ducks and between mallards and domestic ducks in the form of  $D$  (Muscovy duck, Chinese spot-billed duck; P1, P2) and  $D$  (Muscovy duck, mallard; P1, P2), respectively, where P1 and P2 represent different domestic duck groups.

### Demographic history and coalescent simulation

We performed pairwise sequentially Markovian coalescent

(PSMC) analysis (Li & Durbin, 2011) with the setting "--N25 -t15 -r5 -b -p '4+25\*2+4+6'" to infer the demographic history of mallards, Chinese spot-billed ducks, and domestic ducks (highest coverage individual genomes from each group, Supplementary Table S1). A generation time ( $g$ ) of 1 year and a mutation rate per generation of  $\mu=1.91 \times 10^{-9}$  (Zhang et al., 2018) per base pair per year were used for analysis. Multiple sequentially Markovian coalescent (MSMC) analysis (Schiffels & Durbin, 2014) was used to estimate the divergence time between wild and domestic ducks. The ggplot2 R package was used to visualize the PCA plot.

We used fastsimcoal2 (Excoffier et al., 2013) based on joint frequency spectrum data to deduce the recent demographic history of duck populations. EasySFS (<https://github.com/isaacovercast/easySFS#easysfs>) was used to convert variant call format (VCF) to the fastsimcoal2-style site frequency spectrum. Four models were considered, two of which depicted all modern ducks as being domesticated from Chinese spot-billed ducks or mallards, while the other two indicated that domestic ducks originated from a "ghost" duck. Simulations were run with the parameters "--n 100 000 -N 100 000 -d -M 0.001 -l 10 -L40 -q -c 8". For each model, 100 repeat runs with varying starting points were performed to ensure convergence. The best model was selected according to the Akaike information criterion (AIC).

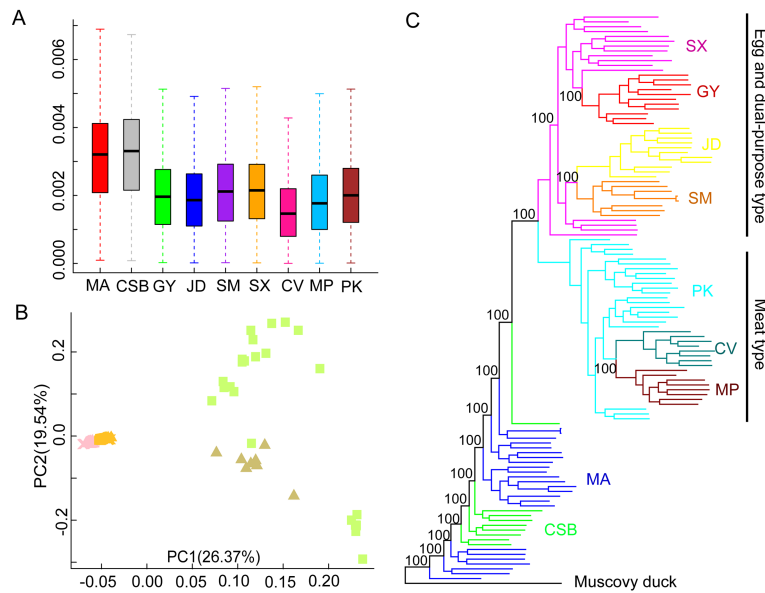
### Simulation analysis of bottleneck effects on MSMC estimation

Recent artificial selection, management, and inbreeding have likely induced bottlenecks during the evolutionary history of domestic animals. Here, we performed simulations to clarify whether bottlenecks have affected our estimation of the relative cross-coalescent rate by MSMC. The software ms (Hudson, 2002) was used, with the command for simulation as follows: /ms 4 100 -t 2292 -r 960 30 000 000 -l 2 2 2 -n 1 X -n 2 2 -en 0.075 1 2 -ej 0.25 1 2 -en 0.25 2 4 -p 7, with  $X=0.25, 0.5, 1, \text{ and } 1.5$  indicating that the current population size is 12.5%, 25%, 50%, and 75% of the ancestral population size ( $2N_0$ ), respectively.

## RESULTS AND DISCUSSION

### Genetic diversity, population relationships, and structure

A total of 118 whole-genome sequences were analyzed in this study. After mapping the reads to the duck reference genome (Zhou et al., 2018), the average sequencing depth for the genomes was estimated at  $\sim 7.5\times$ , ranging from  $3.2\times$  to  $33.9\times$  (Supplementary Table S1). After alignment and quality filtration, more than 14.47 million SNPs were obtained and used for subsequent analysis. We estimated the nucleotide diversity ( $\pi$ ) for each population and found that mallards and Chinese spot-billed ducks had higher levels of nuclear diversity than domestic ducks (Figure 1A). Domestic ducks also showed weaker LD decay than the two wild duck species (Supplementary Figure S2). These results are similar to observations reported in other species such as chickens



**Figure 1 Population relationships and structures of ducks**

A: Population nucleotide diversity. B: Principal component analysis. C: ML tree constructed based on autosomal data. MA: Mallard; CSB: Chinese spot-billed; PK: Pekin; CV: Cherry valley; MP: Maple leaf; GY: Gaoyou; JD: Jinding; SM: Shanma; SX: Shaoxing. Egg and dual-purpose type ducks (EDT), including shanma, shaoxing, jinding, and gaoyou; meat-type ducks (MET), including pekin, cherry valley, and maple leaf.

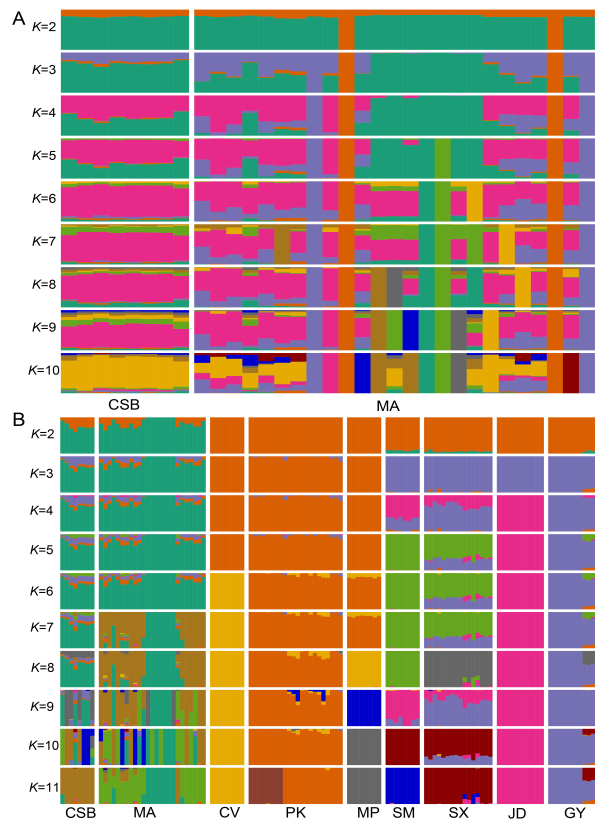
(Wang et al., 2017) and dogs (Wang et al., 2016), suggesting that domestic ducks likely experienced a bottleneck during domestication and management. To clarify the genetic relationships among these populations, we performed PCA (Yang et al., 2011). The main component of genetic variation (PC1) clearly distinguished the wild ducks (Chinese spot-billed ducks and mallards) from the domestic ducks. Within the domestic ducks, EDT and MET were grouped as separate clusters, suggesting potential genetic differentiation within the two populations. We also found that some mallard and Chinese spot-billed duck individuals clustered together (Figure 1B). Population structure analysis using ADMIXTURE (Alexander et al., 2009) showed that Chinese spot-billed ducks and mallards share ancestry (Figure 2A and Supplementary Figures S3, S4), corroborating earlier work based on mtDNA (Wang et al., 2019). Specifically, when  $K=2$ , we found a division between the wild and domestic ducks, whereas the Chinese spot-billed ducks and mallards both harbored a component found predominantly in modern ducks (Figure 2B). When  $K=3$ , we detected a division between EDT and MET. Furthermore, up to  $K=11$ , Chinese spot-billed ducks and mallards presented their own components (Figure 2B). These results suggest that Chinese spot-billed ducks and mallards are closely related.

Both PCA and admixture approaches infer genetic clusters based on data and cannot be used to infer phylogenetic relationships. Therefore, we constructed ML and NJ phylogenetic trees based on whole-genome variants. The mallards and Chinese spot-billed ducks were monophyletic and intermixed, which was strongly supported in both NJ and ML analyses (bootstrap score=100) (Figure 1C and

Supplementary Figure S5). Relationships between domestic ducks were identical in both analyses with high bootstrap support (bootstrap score=100). All domestic ducks clustered together and formed a monophyletic clade sister to wild ducks, consistent with previous research (Zhou et al., 2018), suggesting that modern ducks originated from single domestication. Moreover, we found that the  $F_{ST}$  values for mallards and spot-billed ducks were extremely low ( $\sim 0.01$ , Supplementary Figure S6), suggesting low differentiation. Altogether, these results support a close genetic relationship for the two wild populations, such that it was difficult to distinguish which wild duck group was most closely related to domestic ducks.

#### Introgession between two wild populations and between wild and domestic duck populations

We performed D-statistics (Patterson et al., 2012) in the form of  $D$  (Muscovy duck, mallard, Chinese spot-billed duck, P2) to verify the existence of gene flow between Chinese spot-billed ducks and mallards. When P2 represented Cherry valley, Shanma, Shaoxing, and Gaoyou domestic breeds, it yielded significantly negative Z-scores ranging from  $-17.8$  to  $-6.1$ , indicating that Chinese spot-billed ducks were admixed with mallards (Supplementary Figure S7). Additionally, we also performed D-statistics in the form of  $D$  (Muscovy duck, mallard; P1, P2) and  $D$  (Muscovy duck, Chinese spot-billed duck; P1, P2) to test for admixture between wild and domestic ducks. Our results suggested pervasive introgression patterns between the domestic and wild ducks (Supplementary Table S2). For example, when P3, P2, and P1 represented mallard, Shaoxing, and other domestic ducks, respectively,



**Figure 2** Admixture analysis of mallards and Chinese spot-billed ducks (A) and mallards, Chinese spot-billed ducks, and domestic ducks (B)

MA: Mallard; CSB: Chinese spot-billed; PK: Pekin; CV: Cherry valley; MP: Maple leaf; GY: Gaoyou; JD: Jinding; SM: Shanma; SX: Shaoxing.

they yielded scores ranging from  $-20.1$  to  $-4.6$ . These results are not surprising as hybridization is common in birds (Grant & Grant, 1992), particularly in Anatidae (Lavretsky et al., 2020; Muñoz-Fuentes et al., 2007; Peters et al., 2007). Mallards have a wide distribution, spanning the British Isles through Europe and Russia to northern Japan and northeastern China, while spot-billed ducks are abundant in northern and western China, the Korean Peninsula, and Japan (Kulikova et al., 2003). In some regions, their distribution range overlaps and hybridization between the two species has been pervasive and frequent. During the last decade, spot-billed duck and mallard hybridization has been reported in Khanka Lake and Hong Kong (Kulikova et al., 2003). The complex introgression patterns among these duck populations may confound domestic duck origin and evolutionary history results when using mtDNA and short nuclear segments.

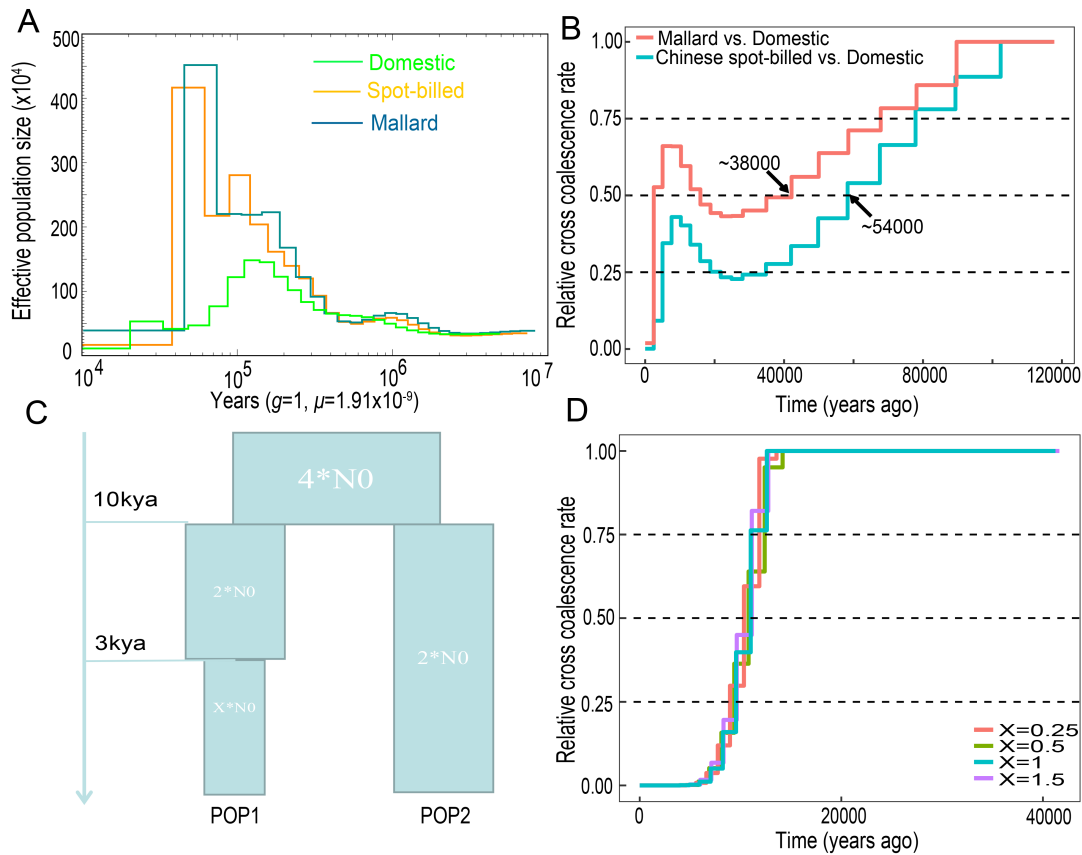
#### Demographic history and divergence time

If domestic ducks originated from one of the two wild duck species, we would expect them to have a demographic history similar to that of wild ducks before the Holocene (Larson et al.,

2014). We used PSMC analysis (Li & Durbin, 2011) to estimate the evolution of historical effective population sizes ( $N_e$ ) of the wild and domestic ducks. Domestic ducks were found to have undergone a moderate expansion of  $N_e$  beginning approximately one million years ago (mya), which peaked  $\sim 120$  kya and then declined for an estimated period of tens of thousands of years (Figure 3A). The two wild duck populations showed similar  $N_e$  trajectories. The  $N_e$  for Chinese spot-billed ducks and mallards increased  $\sim 1$  mya and reached a maximum  $\sim 40$  kya, after which it contracted dramatically (Figure 3A). These  $N_e$  fluctuations are similar to those observed in chickens and ostriches (Nadachowska-Brzyska et al., 2015; Wang et al., 2015). During the Mid-Brunhes event (430–110 kya), an interglacial period in which the climate was warm, many species showed population size expansion. However, when entering the last ice age (110–10 kya), global temperatures declined drastically (Kozma et al., 2016), which induced loss of habitat and distribution contraction for many species, resulting in reductions in population size (Nadachowska-Brzyska et al., 2015). Therefore, we suggest that, similar to other birds, climate change during the most recent glaciations likely affected the  $N_e$  of ducks.

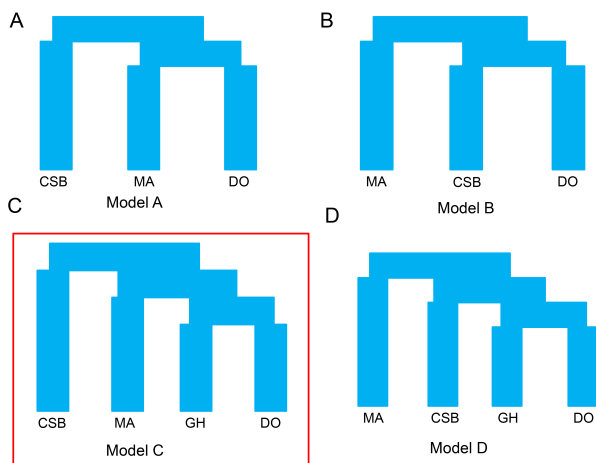
To our surprise, the evolutionary trajectories of the two wild duck populations were found to be distinct from those of domestic ducks beginning 100 kya (Figure 3A). We used MSMC analysis (Schiffels & Durbin, 2014) to measure the divergence time between the wild and domestic ducks. Results showed that domestic ducks separated from mallards and Chinese spot-billed ducks nearly 38 kya and 54 kya (based on  $CCR=0.5$ ) (Figure 3B), respectively, which are largely outside the time period of animal domestication and presumed duck domestication (Hitosugi et al., 2007; Larson et al., 2014). However, MSMC estimation mainly depends on the number and density of heterozygotes, and strong selection in modern ducks may have induced the accumulation of homozygotes and a reduction in heterozygotes, thereby confounding our estimation. To test this, we performed simulations using a model with two populations (split 10 kya), in which one experienced multiple bottleneck events beginning 3 kya (Hitosugi et al., 2007). The obtained estimates indicated that our MSMC analysis was sufficient for estimating divergence (Figure 3C, D). These results suggested that Chinese spot-billed ducks and mallards are unlikely to be the wild relative from which domestic ducks descended. Thus, we used the joint site frequency spectrum (SFS) approach implemented in fastsimcoal2 to compare four scenarios. The simulation results also supported the origination of domestic ducks from a ghost duck species (Figure 4 and Supplementary Figure S8). These findings imply that both Chinese spot-billed ducks and mallards are unlikely to be the direct wild ancestor of modern ducks.

As wild mallards and spot-billed ducks are widely distributed, they habitually undergo long-distance annual



**Figure 3 Demographic history of ducks**

A: Demographic history inferred by PSMC with a generation time=1 year and mutation rate= $1.91 \times 10^{-9}$ . B: Inferred relative cross-coalescence rates between pairs of populations over time based on two individuals of spot-billed, mallard, and domestic ducks. C: Demographic scenarios simulated by ms software ( $X=0.25, 0.5, 1, \text{ and } 1.5$ , representing current population sizes equal to 12.5%, 25%, 50%, and 75% of ancestral population size, respectively). D: MSMC analysis of simulated data.



**Figure 4 Demographic scenarios tested by fastsimcoal2**

A, B: Ducks domesticated from Chinese spot-billed ducks or mallards. C, D: Ducks domesticated from "ghost" duck population. MA: Mallard; CSB: Chinese spot-billed; GH: Ghost; DO: Domestic. Model C is best.

migrations (Kulikova et al., 2005; Shin et al., 2015; Williams et al., 2020). Therefore, it is plausible that there are some substructure wild duck populations that have not been sampled or have gone extinct. Domestic ducks likely descended from such unknown wild duck populations and not from present-day Chinese mallards or spot-billed ducks, as previously assumed (Hou et al., 2012). Interestingly, our results are similar to observations for other domestic species. For example, genomic analysis has shown that dogs originated from an extinct wolf population (Fan et al. 2016) and domestic horses were derived from an extinct horse population (Fages et al., 2019; Gaunitz et al., 2018). Our study suggests that the evolutionary history of ducks is more complex than previously thought and provides a possibility for future exploration. Further studies incorporating additional wild, domestic, and ancient samples covering all duck domestication sites are necessary to test our hypothesis and elucidate the evolutionary history of this important poultry species.

## CONCLUSIONS

Our results provide consistent and compelling evidence that Chinese spot-billed ducks and mallards are genetically closely related. Moreover, we also found a pervasive and complex introgression pattern between wild and domestic ducks. In particular, our results suggest that domestic ducks are unlikely to have originated from present-day mallards and Chinese spot-billed ducks, as previously assumed, but may have originated from a ghost wild duck population that is likely undefined and/or not included in the present study. Overall, our study provides new insight into the complex evolution of ducks and provides a scenario for further study.

## SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

## COMPETING INTERESTS

The authors declare that they have no competing interests.

## AUTHORS' CONTRIBUTIONS

R.S.J. and M.S.W. designed the study and revised the paper; X.G., H.C., Z.C.W., and H.F.L. conducted bioinformatic analyses; X.G. wrote the draft manuscript. All authors read and approved the final version of the manuscript.

## REFERENCES

- Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, **19**(9): 1655–1664.
- Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, **4**(1): 7.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**(15): 2156–2158.
- Dong Y, Xie M, Jiang Y, Xiao NQ, Du XY, Zhang WG, et al. 2013. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nature Biotechnology*, **31**(2): 135–141.
- Excoffier L, Dupanloup I, Huerta-Sanchez E, Sousa VC, Foll M. 2013. Robust demographic inference from genomic and SNP data. *PLoS Genetics*, **9**(10): e1003905.
- Fages A, Hanghøj K, Khan N, Gaunitz C, Seguin-Orlando A, Leonardi M, et al. 2019. Tracking five millennia of horse management with extensive ancient genome time series. *Cell*, **177**(6): 1419–1435.e31.
- Fan ZX, Silva P, Gronau I, Wang SG, Armero AS, Schweizer RM, et al. 2016. Worldwide patterns of genomic variation and admixture in gray wolves. *Genome Research*, **26**(2): 163–173.
- Gaunitz C, Fages A, Hanghøj K, Albrechtsen A, Khan N, Schubert M, et al. 2018. Ancient genomes revisit the ancestry of domestic and Przewalski's horses. *Science*, **360**(6348): 111–114.
- Grant PR, Grant BR. 1992. Hybridization of bird species. *Science*, **256**(5054): 193–197.
- Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, et al. 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, **491**(7424): 393–398.
- Guo X, Fang Q, Ma CD, Zhou BY, Wan Y, Jiang RS. 2016. Whole-genome resequencing of Xishuangbanna fighting chicken to identify signatures of selection. *Genetics Selection Evolution*, **48**(1): 62.
- Hillier LDW, Miller W, Birney E, Warren W, Hardison RC, Ponting CP, et al. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, **432**(7018): 695–716.
- Hitosugi S, Tsuda K, Okabayashi H, Tanabe Y. 2007. Phylogenetic relationships of mitochondrial DNA cytochrome b gene in east asian ducks. *The Journal of Poultry Science*, **44**(2): 141–145.
- Hou ZC, Yang FX, Qu LJ, Zheng JX, Brun JM, Basso B, et al. 2012. Genetic structure of Eurasian and North American mallard ducks based on mtDNA data. *Animal Genetics*, **43**(3): 352–355.
- Huang YH, Li YR, Burt DW, Chen HL, Zhang Y, Qian WB, et al. 2013. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nature Genetics*, **45**(7): 776–783.
- Hudson RR. 2002. Generating samples under a wright-fisher neutral model of genetic variation. *Bioinformatics*, **18**(2): 337–338.
- Jin SD, Hoque R, Seo DW, Paek WK, Kang TH, Kim HK, et al. 2014. Phylogenetic analysis between domestic and wild duck species in Korea using mtDNA D-loop sequences. *Molecular Biology Reports*, **41**(3): 1645–1652.
- Kong Y. 2011. Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics*, **98**(2): 152–153.
- Kozma R, Melsted P, Magnusson KP, Höglund J. 2016. Looking into the past - the reaction of three grouse species to climate change over the last million years using whole genome sequences. *Molecular Ecology*, **25**(2): 570–580.
- Kulikova IV, Chelomina GN, Zhuravlev YN. 2003. Low genetic differentiation of and close evolutionary relationships between *Anas platyrhynchos* and *Anas poecilorhyncha*: RAPD-PCR evidence. *Russian Journal of Genetics*, **39**(10): 1143–1151.
- Kulikova IV, Drovetski SV, Gibson DD, Harrigan RJ, Rohwer S, Sorenson MD, et al. 2005. Phylogeography of the mallard (*Anas platyrhynchos*): hybridization, dispersal, and lineage sorting contribute to complex geographic structure. *The Auk*, **122**(3): 949–965.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7. 0 for bigger datasets. *Molecular Biology and Evolution*, **33**(7): 1870–1874.
- Larson G, Burger J. 2013. A population genetics view of animal domestication. *Trends in Genetics*, **29**(4): 197–205.
- Larson G, Piperno DR, Allaby RG, Purugganan MD, Andersson L, Arroyo-Kalin M, et al. 2014. Current perspectives and the future of domestication studies. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(17): 6139–6146.
- Lavretsky P, McCracken KG, Peters JL. 2014. Phylogenetics of a recent radiation in the mallards and allies (Aves: Anas): inferences from a genomic transect and the multispecies coalescent. *Molecular Phylogenetics and Evolution*, **70**: 402–411.
- Lavretsky P, McInerney NR, Mohl JE, Brown JI, James HF, McCracken KG, et al. 2020. Assessing changes in genomic divergence following a century of human-mediated secondary contact among wild and captive-bred ducks.

- Molecular Ecology*, **29**(3): 578–595.
- Li H. 2014. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics*, **30**(20): 2843–2851.
- Li H, Durbin R. 2011. Inference of human population history from individual whole-genome sequences. *Nature*, **475**(7357): 493–496.
- Li HF, Zhu WQ, Song WT, Shu JT, Han W, Chen KW. 2010. Origin and genetic diversity of Chinese domestic ducks. *Molecular Phylogenetics and Evolution*, **57**(2): 634–640.
- Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, et al. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, **438**(7069): 803–819.
- Liu YP, Wu GS, Yao YG, Miao YW, Luikart G, Baig M, et al. 2006. Multiple maternal origins of chickens: out of the Asian jungles. *Molecular Phylogenetics and Evolution*, **38**(1): 12–19.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**(9): 1297–1303.
- Miao YW, Peng MS, Wu GS, Ouyang YN, Yang ZY, Yu N, et al. 2013. Chicken domestication: an updated perspective based on mitochondrial genomes. *Heredity*, **110**(3): 277–282.
- Muñoz-Fuentes V, Vilà C, Green AJ, Negro JJ, Sorenson MD. 2007. Hybridization between white-headed ducks and introduced ruddy ducks in Spain. *Molecular Ecology*, **16**(3): 629–638.
- Museum ZC. 1979. The excavation of the mound tomb at Guoyuan of Fushan, Jurong County, Jiangsu Province. *Kaogu*, **1979**(2): 113.
- Nadachowska-Brzyska K, Li C, Smeds L, Zhang GJ, Ellegren H. 2015. Temporal dynamics of avian populations during pleistocene revealed by whole-genome sequences. *Current Biology*, **25**(10): 1375–1380.
- Orlando L, Ginolhac A, Zhang GJ, Froese D, Albrechtsen A, Stiller M, et al. 2013. Recalibrating *Equus* evolution using the genome sequence of an early middle pleistocene horse. *Nature*, **499**(7456): 74–78.
- Park SDE, Magee DA, McGettigan PA, Teasdale MD, Edwards CJ, Lohan AJ, et al. 2015. Genome sequencing of the extinct Eurasian wild aurochs, *Bos primigenius*, illuminates the phylogeography and evolution of cattle. *Genome Biology*, **16**(1): 234.
- Patterson N, Moorjani P, Luo YT, Mallick S, Rohland N, Zhan YP, et al. 2012. Ancient admixture in human history. *Genetics*, **192**(3): 1065–1093.
- Peters JL, Zhuravlev Y, Fefelov I, Logie A, Omland KE. 2007. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas* spp.). *Evolution*, **61**(8): 1992–2006.
- Plassais J, Kim J, Davis BW, Karyadi DM, Hogan AN, Harris AC, et al. 2019. Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nature Communications*, **10**(1): 1489.
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS One*, **5**(3): e9490.
- Savolainen P, Zhang YP, Luo J, Lundeberg J, Leitner T. 2002. Genetic evidence for an East Asian origin of domestic dogs. *Science*, **298**(5598): 1610–1613.
- Schiffels S, Durbin R. 2014. Inferring human population size and separation history from multiple genome sequences. *Nature Genetics*, **46**(8): 919–925.
- Shin JH, Lee KS, Kim SH, Hwang JK, Woo C, Kim J, et al. 2015. Tracking mallards (*Anas platyrhynchos*) with GPS satellite transmitters along their migration route through Northeast Asia. *Avian Diseases*, **60**(1S): 311–315.
- Thalmann O, Shapiro B, Cui P, Schuenemann VJ, Sawyer SK, Greenfield DL, et al. 2013. Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science*, **342**(6160): 871–874.
- Wang GD, Xie HB, Peng MS, Irwin D, Zhang YP. 2014. Domestication genomics: evidence from animals. *Annual Review of Animal Biosciences*, **2**: 65–84.
- Wang GD, Zhai WW, Yang HC, Wang L, Zhong L, Liu YH, et al. 2016. Out of southern East Asia: the natural history of domestic dogs across the world. *Cell Research*, **26**(1): 21–33.
- Wang MS, Li Y, Peng MS, Zhong L, Wang ZJ, Li QY, et al. 2015. Genomic analyses reveal potential independent adaptation to high altitude in Tibetan chickens. *Molecular Biology and Evolution*, **32**(7): 1880–1889.
- Wang MS, Otecko NO, Wang S, Wu DD, Yang MM, Xu YL, et al. 2017. An evolutionary genomic perspective on the breeding of dwarf chickens. *Molecular Biology and Evolution*, **34**(12): 3081–3088.
- Wang MS, Wang S, Li Y, Jhala Y, Thakur M, Otecko NO, et al. 2020. Ancient hybridization with an unknown population facilitated high-altitude adaptation of canids. *Molecular Biology and Evolution*, **37**(9): 2616–2629.
- Wang WJ, Wang YF, Lei FM, Liu Y, Wang HT, Chen JK. 2019. Incomplete lineage sorting and introgression in the diversification of chinese spot-billed ducks and mallards. *Current Zoology*, **65**(5): 589–597.
- Williams BR, Benson TJ, Yetter AP, Lancaster JD, Hagy HM. 2020. Habitat use of spring migrating dabbling ducks in the Wabash River Valley, USA. *The Condor*, **122**(1): duz061.
- Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, **88**(1): 76–82.
- Zhang Y, Chen Y, Zhen T, Huang ZY, Chen CY, Li XY, et al. 2014. Analysis of the genetic diversity and origin of some chinese domestic duck breeds. *Journal of Integrative Agriculture*, **13**(4): 849–857.
- Zhang Z, Khederzadeh S, Li Y. 2020. Deciphering the puzzles of dog domestication. *Zoological Research*, **41**(2): 97–104.
- Zhang ZB, Jia YX, Almeida P, Mank JE, van Tuinen M, Wang Q, et al. 2018. Whole-genome resequencing reveals signatures of selection and timing of duck domestication. *Gigascience*, **7**(4): gjy027.
- Zhou ZK, Li M, Cheng H, Fan WL, Yuan ZR, Gao Q, et al. 2018. An intercross population study reveals genes associated with body size and plumage color in ducks. *Nature Communications*, **9**(1): 2648.